Introduction
**Introduction**

The work carried out in the present dissertation deals with the biochemical characterization of the nucleoid from the thermoacidophilic crenarchaeote *Sulfolobus acidocaldarius* and the helix stabilizing nucleoid protein, HSNP A.

A brief account of the archaea, their molecular biological properties, the chromatin organization in archaea, the proteins associated with archaeal nucleoid, the cell cycle and the associated biochemical variations observed is given in the following pages.

**Archaea: a brief introduction:**

Domain archaea comprise an evolutionary lineage of prokaryotes only distantly related to bacteria (Woese and Fox, 1977; Woese et al., 1978; Woese et al., 1990). Archaea from geo and hydrothermal habitats include the most thermophilic organisms known (Stetter, 1999), they represent the shortest and deepest branches in the universal tree of life and exhibit features formerly thought to be uniquely eukaryotic. Therefore studies of thermophilic archaea provide insights into mechanisms of thermal stability and the fundamental evolutionary relationships among cellular organisms.

Archaeal features can be classified as bacteria-like, eukaryotic-like or unique to the archaea. As much as a third of the archael genome consists of genes found uniquely in archaea (Woese and Gupta, 1981).

**General features of the Archaea**

**Morphology:** Phenotypically, the Archaea are a more like Bacteria. Most are small (0.5-5 microns) rods, cocci, spirilla, and filaments. Many of them grow as clusters or aggregates. Archaea generally reproduce by fission, like Bacteria and most Eukarya. The genomes of Archaea are generally 2- 4Mbp in size, similar to Bacteria. However, majority of Archaea are thermophilic - in fact, many are *extremely* thermophilic. They
are also either autotrophic or sulfur respirers. Like eukaryotes, some Archaea have abundant histone-like proteins and the DNA is packaged in the form of nucleosomes.

**Cell envelopes:** The Archaeal cells are generally surrounded by complex and diverse cell envelopes with glycoprotein containing S-layers and cytoplasmic membranes containing tetraether lipids. Archaeal membrane lipids are unique ether-linked (not ester-linked) glycerol derivatives comprising a glycerol diether with two C$_{20}$ chains and diglycerol tetra ether with two C$_{40}$ chains. Unsaturations in the lipid chain are generally conjugated (those of Bacteria & eukaryotes are unconjugated). 40-carbon lipids are ether-linked to glycerol at ends, and so form lipid monolayers rather than bilayers, with polar groups at each end. This unique structure increases the chemical and thermal stability of the membrane and may reduce excess fluidity of the membrane at extremely high temperatures. Archaea are incapable of adapting to temperature changes by altering the degree of unsaturation or chain length of membrane lipids like the mesophiles. Increase in growth temperature has been shown to increase the frequency of cyclopentane rings in *Sulfolobus solfataricus* and *Thermoplasma acidophilum*. These lipids can be used as chemical signatures for the identification of Archaea in a sample (Kandler and König, 1985).

Archaea are further classified into three sub domains, the Euryarchaeota, the Crenarchaeota and the Korarchaeota. The Korarchaeota have been deduced from small subunit rDNA sequence data, they have not been isolated or cultivated as yet. (Barns et al., 1996).

Members of the Euryarchaeota include the methanogens and halophiles. Methanogenic archaea use three compounds as carbon sources, methyl alcohol, acetate and CO$_2$ -type substrates and produce methane by reducing these compounds and oxidizing H$_2$. Methanogens cannot tolerate O$_2$ and are thus found in anoxic environments like swamps.

The halophilic archaea include aerobes. These organisms are found in salty or alkaline environments, some require saturated salt solutions for growth. All of the extreme halophilic archaea are chemoorganotrophic.
Some members of the Euryarchaeota are hyperthermophiles and can grow in extremely acidic environments (<pH 2.0).

Members of the Korarchaeota were initially based on recovered rDNA sequences from volcanic hot springs. They have been detected in mixed cultures using fluorescent DNA probes. These organisms probably have similar metabolisms to other Archaea from extreme environments. The Korarchaeota branch near the division between bacteria, archaea and eukaryotes.

Cultivated crenarchaeota are all thermophilic, and most are extremely thermophilic, with optimal growth temperatures above 80 °C. As a group, these are the most thermophilic organisms known that grow between 80° and 100°C; the maximum temperature of growth can be as high as 113°C (Pyrobolus). Most of the hyperthermophilic species have been isolated from hot springs and shallow or deep-sea hydrothermal vents. Some species can live in dilute solutions of hot sulfuric acid; members of the order Sulfolobales grow at pH 1-2. Low temperatures Crenarchaeota have been identified by genetic analysis of environmental samples. rDNA signatures of Crenarchaeota species have been detected in low-temperature environments such as ocean waters (Antarctic Ocean) and terrestrial sediments and soils.

The Crenarchaeota are more diverse both phenotypically and phylogenetically than the cultivated species would imply. Analysis of a single sample from Obsidian Pool in Yellowstone revealed the presence of more crenarchaeal species with more phylogenetic range than all of the cultivated Crenarchaeota taken together. Crenarchaeal small subunit-rRNA sequences unrelated to cultivated species are commonly isolated from forest soil and lake sediment, suggesting the existence of species that grow in moderate conditions. One species, Crenarchaeum symbiosum, which so far has not been grown in culture, is a symbiont of marine sponges.

Several novel methods currently in use to detect previously unidentified microbes show archaea to occur widely in nature. Novel archaea have been localized in agricultural (Kudo et al., 1997, Bintrim et al., 1997) and forest soils (Jurgens et al., 1997) in fresh water lake sediments (Schleper, et al., 1997), in coastal water.
<table>
<thead>
<tr>
<th>Attribute</th>
<th>Bacteria</th>
<th>Archaea</th>
<th>Eukaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell structure</td>
<td>prokaryotic</td>
<td>prokaryotic</td>
<td>eukaryotic</td>
</tr>
<tr>
<td>DNA</td>
<td>mostly circular</td>
<td>mostly circular</td>
<td>linear chromosomes</td>
</tr>
<tr>
<td>Histones present</td>
<td>No</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>cell wall</td>
<td>peptidoglycan (muramic acid)</td>
<td>muramic acid absent</td>
<td>muramic acid absent</td>
</tr>
<tr>
<td>membrane lipids</td>
<td>ester-linked</td>
<td>ether-linked</td>
<td>ester-linked</td>
</tr>
<tr>
<td>Ribosomes</td>
<td>70S</td>
<td>70S</td>
<td>80s</td>
</tr>
<tr>
<td>Initiator tRNA</td>
<td>N-formylmethionine</td>
<td>methionine</td>
<td>methionine</td>
</tr>
<tr>
<td>Introns</td>
<td>no (mostly)</td>
<td>no (mostly)</td>
<td>yes (mostly)</td>
</tr>
<tr>
<td>Operons</td>
<td>Yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Plasmids</td>
<td>common</td>
<td>common</td>
<td>rare</td>
</tr>
<tr>
<td>RNA polymerase</td>
<td>one (4 subunits)</td>
<td>One (8-12 subunits)</td>
<td>three (12-14 subunits each)</td>
</tr>
<tr>
<td>Promoter structure</td>
<td>-10 and -35 (Pribnow box)</td>
<td>TATA box</td>
<td>TATA box</td>
</tr>
<tr>
<td>Sensitivity to chloramphenicol, kanamycin and streptomycin</td>
<td>Yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Methanogenesis</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Nitrification</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Nitrogen fixation</td>
<td>yes</td>
<td>yes</td>
<td>yes (chloroplasts)</td>
</tr>
<tr>
<td>Chlorophyll-based photosynthesis</td>
<td>yes</td>
<td>no</td>
<td>yes (chloroplasts)</td>
</tr>
<tr>
<td>Chemolithotrophy</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Growth above 80C</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>
picoplankton (Delong, 1994), in deep waters up to 3000 m of the oceans (Fuhrman et al., 1997) besides the well established extreme environments.

Crenarchaeota metabolism ranges from chemoorganotrophy to *chemolithoautotrophy*. Most are also acidophilic and *autotrophic*. Since this phenotype is shared by the most primitive & first branches of the Euryarchaea, and even the Bacteria, it is probably the primitive phenotype of the Archaea as well as the last common ancestor.

The Crenarchaea fall into 2 subgroups: *Thermoproteus* and relatives (7 genera) and *Sulfolobus* and relatives (5 genera).

**Thermoproteus and related species:** They are strict anaerobes that grow best by sulfur respiration under culture conditions, the more usual growth conditions for these organisms in the wild is autotrophic sulfur reduction i.e., these organisms can either oxidize or reduce sulfur and the switch between growth modes is a distinct developmental switch. These organisms are anaerobic, extreme thermophiles. *Pyrodictium spp* cultures grow optimally at up to 115 °C and are among the most primitive organisms known. Related species are *Desulfurococcus, Staphylothermus*, and *Thermodiscus*.

**Sulfolobus and related species:** These organisms are common, generally predominant, organisms of solfataras, boiling mud pots, and hot acid mine drainage. They are generally facultative *microaerophiles*, growing best in culture by sulfur oxidation (aerobically or microaerophilically) or sulfur reduction (microaerophilically or anaerobically). They are all acidophilic, growing best at pH 1 to 3.5, although a couple of species grow at pH 0. They are also all *thermophilic*, growing optimally at 75 to 95 °C. These organisms like most acidophiles, are oligotrophic and require small amounts of organics for growth. Organic acids in particular are toxic to acidophiles since they are protonated at the pH's of the growth media (below 3.5), and are therefore uncharged, facilitating them to diffuse freely into the cytoplasm through the lipid membrane. At pH >5.5, the internal pH of the cells, the organic acids ionize, releasing H+. This both acidifies the cytoplasm and decomposes the proton gradient which is harmful to the cell. Related genera are *Metallosphaera, Acidianus, Desulfurolobus*, and *Stygiolobus*.  

5
**Sulfolobus acidocaldarius**

*Sulfolobus acidocaldarius* grows optimally at 80°C and pH 3 and it is a lobed coccus. The lobes are budding scars, left over from budding reproduction, and appear like appendages that hold the cell to the substrate sulfur granules. *Sulfolobus* is a strict microaerophile, first discovered in Moose pool (Yellowstone National Park) where it grows nearly in pure culture at about $10^8$ cells per ml. *Sulfolobus spp* have developed as an important model for studies of hyperthermophilic archaea, including cell cycle analyses (Bernander and Poplawski, 1997; Hjort and Bernander, 1999; Poplawski and Bernander, 1997) and development of simple genetic techniques (Grogan, 1989, 1996; Grogan and Gunsalus, 1993; Schleper et al., 1994).

Studies on the cell cycle of *Sulfolobus* demonstrate that the chromosome replication, cell division, nucleoid structure and distribution and cellular growth can be selectively inhibited or uncoupled (Bernander et al., 2000). The post replication stage has been determined to be the preferred cell cycle arrest stage. A method for synchronizing the *Sulfolobus acidocaldarius* cultures has been developed by Hjort and Brenander, 1999. When a stationary phase culture, in which all cells are in the post replication stage of the cell cycle, is diluted into fresh medium, initiation of replication does not occur until the preceding nucleoid partition and cell division.

No analogue of the *ftsZ* (filamentation temperature sensitive) gene, whose product initiates bacterial cell division and participates in the constriction process, has been identified in a crenarchaeote.

Archaea differ considerably in their biology from Bacteria and Eukaryotes - at least 30% of their genes have no detectable homologues in the other two domains - they are a rich source of novel enzymes. Moreover, since many of the archaeal homologs of eukaryotic *multi-subunit* enzymes are *monomeric*, this greatly extends their applicability to biotechnological processes.

In 2001, The European *Sulfolobus* Project was initiated to screen for and express thermostable enzymes of industrial interest from *Sulfolobus* and other Crenarchaeotes. The purpose of this project is to exploit the expertise of the european laboratories.
involved in studies on this organism and to screen for novel enzyme activities both at the genomic level and in an extensive and unique European collection of thermophilic Crenarchaeotes, to develop and refine shuttle vector systems for expressing these proteins in a range of Bacterial and Eukaryotic hosts by modifying the Archaeal genetic signals, to develop a vector-host system for expressing Crenarchaeal enzymes in a *Sulfolobus* host on a large scale, to perform large-scale fermentation to express new crenarchaeal enzymes and to thermoadapt mesophilic enzymes in *Sulfolobus*. The project involves a genomic approach where we exploit the finished genome of *S. solfataricus* (She et al., 2001) and perform low pass sequencing of the genomes of two other diverse and divergent Crenarchaeotes *Sulfolobus acidocaldarius* and *Hyperthermus butylicus*.

**Molecular biology of Archaea:**

Among the major cellular functions to be characterized in archaea are chromosome replication, nucleoid segregation and cell division, and the coordination of these processes with cellular growth. Analysis of the archaeal genome sequences has revealed the replication, transcription and translation machineries of archaea to be similar to those of eukaryotic organisms, while genes involved in the metabolic processes show more similarity to their bacterial counterparts. For example, the gene products involved in the chromosome segregation show similarity to both eukaryal and bacterial proteins whereas the cell division apparatus is mostly of bacterial type. (Bernander, 1998). Since the archaea cell cycle contains a mixture of features from both eukarya and bacteria they may contribute to a better understanding of the evolution of these processes. There is speculation that the eukaryotes may have originated from a fusion event between a bacterium and an archaeon (Gupta and Golding, 1996; Martin and Muller, 1998; Moreira and Lopez-Garcia, 1998) which implies that the eukaryal cell cycle characteristics may be derived from those of an archaea like ancestor.

Certain features however are characteristic of archaea like their ribosomal RNA sequences, the lipid composition (ether glycerol lipids) and organization (bi or mono layer) of the cytoplasmic membrane and also particular modes of energy metabolism,
such as methanogenesis and bacterio rhodopsin- driven photosynthesis (Jones et al., 1987)

**DNA replication machinery:** *Sulfolobus solfataricus* has two functional DNA polymerases, PolB1 and PolY1. PolB1 specifically recognizes the presence of deaminated bases hypoxanthine and uracil in the template by stalling polymerization 3-4 bases upstream of these lesions and strongly associates with oligonucleotides containing them (Greagg et al., 1999). PolB1 also stops at 8-oxoguanine and is unable to bypass an abasic site in the template. PolY1 belongs to the family of lesion bypass DNA polymerases and readily bypasses hypoxanthine, uracil and 8-oxoguanine, but not an abasic site in the template. Both PolB1 and PolY1 incorporate aberrant DNA precursors dUTP and dITP (Gruz et al., 2003). The error prone PolY1 a homolog of DinB uses a steric gate residue for discriminating against incorporation of ribonucleotides during polymerization. (DeLucia et al., 2003).

An 800 bp intergeneric sequence between the genes encoding Cdc6/Orc1* and methylthioadenosine phosphorylase which is highly conserved in the *Pyrococci* species has been identified as the origin for replication in this organism (Kelman, 2000). Two active regions of replication have been identified in the chromosome of the hyperthermophilic archaean *Sulfolobus solfataricus*. Conserved sequences are present within these regions which are recognized by a family of three *Sulfolobus* proteins homologous to Orc 1 and Cdc6 proteins. The two regions are recognized by distinct subsets of these Orc1/Cdc6 homologs (Robinson et al., 2004).

PCNA (proliferating cell nuclear antigen) has been shown to be functionally conserved not only within eukaryotes but also from hyperthermophilic euryarchaeotes to mammals (Henneke et al., 2000). The sliding clamp, PCNA of the archaean *Sulfolobus solfataricus* P2 is a heterotrimer of three distinct subunits (PCNA1, 2, 3) that assembles in a defined manner. The PCNA heterotrimer, but not the individual subunits, stimulates the activities of DNA polymerase, DNA ligase I and the flap endonuclease (FEN1) of *S.solfataricus*. Unique subunit specific interactions between components of the clamp loader, RFC, suggest a model for clamp loading of PCNA (Dionne, I., et al., 2003).
Primase activity of eukaryotes resides within the DNA polymerase α-primase complex (pol α-prim) on the p48 subunit. A p48-like DNA primase from *Pyrococcus furiosus*, *Pfup41*, preferentially utilizes deoxynucleotides to synthesize DNA fragments up to several kilobases in length without primers. Reports that a DNA polymerase can be engineered from an RNA polymerase by site specific mutagenesis support the idea that DNA polymerases might have evolved from RNA polymerases. *Pfup41* may be an intermediate in such pathways between an RNA polymerase-primase and a DNA polymerase (Bocquier et al., 2001).

Archaeal minichromosome maintenance (MCM) helicase moves along ds DNA even in the absence of a 3' ss DNA overhang (Shin et al., 2003). DNA topology of hyperthermophilic archaea during stress relies on the physical effect of temperature on topoisomerase activity and on the DNA geometry (Lopez-Garcia et al., 1999). Sso MCM has NTPase and helicase activity which acts preferentially on DNA duplexes containing a 5' tail and is stimulated by the single stranded DNA binding protein SSB. (Carpentieri, et al., 2002)

The first member of the DNA polymerase E family has been reported from *Sulfolobus islandicus* (Lipps et al., 2003). ORF904, encoded by a plasmid pRN1 has ATPase, primase and DNA polymerase activity. The primase and polymerase activity are located in the N-terminal half of the protein, which does not share homology with any known DNA polymerase or primase. The ORF904 type of replication enzyme could have evolved independently from the eubacterial and archaeal/eukaryal proteins of DNA replication.

**Transcription and translation machinery:** The transcriptional/translational machinery in Archaea is generally like those of Bacteria and Eukarya, with 70S (bacterial-sized) ribosomes. Genes are arranged in co-transcribed clusters called operons. Ribosomes recognize translational start sites and bind to the mRNAs directly at 'Shine-Dalgarno' (SD) sequences just like Bacteria. Also like in Bacteria, transcription and translation are coupled - that is, they occur simultaneously, and failure of an mRNA to be translated causes the RNA polymerase to abort transcription. Introns in protein coding genes are ubiquitous in eukaryotes but are
rarely found in prokaryotes. Recent evidence through computational analysis has revealed the presence of introns in protein encoding gene, the Cbf5p (centeromere-binding factor 5) homolog, in the genomes of the archaea, *Sulfolobus solfataricus*, *S. tokodaii* and *Aeropyrum pernix* (Watanabe et al., 2002).

However, in many ways transcription and translation in Archaea is like it is in Eukarya. Although each promoter drives the expression of an entire operon, the promoters are just like eukaryotic RNA polymerase II promoters, and are binding sites for transcription factors (rather than the RNA polymerase itself, as in Bacteria). Archaea have a single RNA polymerase (like Bacteria), but this RNA polymerase is just like eukaryotic RNA polymerase II, and requires the same transcription factors for promoter recognition. Translation is initiated with methionine (like Eukarya), not formyl-methionine (Bacteria). Translation is inhibited by diphtheria toxin, as eukaryal ribosomes, but is not inhibited by most bacterial-translation-inhibiting antibiotics. Hybrid archaeal/eukarya ribosomes are functional where as bacterial/archaeal or bacterial /eukarya ribosomes are not functional.

**Cell cycle characteristics of Archaea:**

The cell cycle characteristics of archaea share similarities with both prokarya and eukarya. Growth phase dependent modifications of nucleoid size and a long gap period between the termination of chromosome replication and completion of nucleoid partition, similar to the G2 phase in eukaryotic cells were observed in the euryarchaeon *Thermococcus kodakaraensis* (Jeon., et al., 2001) and the crenarchaeon *Sulfolobus acidocaldarius* (Poplawski and Bernander., 1997). The period of visible cell constriction, however was found to be similar to that of prokaryotes. The separation of nucleoids to opposite cell halves before division is similar to a recent model proposed for nucleoid partition in eubacteria (Wheeler and Shapiro., 1997) in which the replication origins become attached to the cell poles before division. During the exponential phase the nucleoids are highly organized but have an unstructured appearance during stationary phase. The structure of *E. coli* nucleoid has been suggested to be highly dependent on the transcriptional activity through coupled transcription and translation of membrane proteins which would anchor transcription
complexes and the associated DNA in the membrane and thereby extend the nucleoid. (Woldringh and Nanninga, 1985; Woldringh et al., 1995). If coupled transcription, translation and membrane insertion also occurs in archaea, the relaxed nucleoid structure in stationary phase may at least in part due to reduced transcriptional activity.

Genome projects have revealed the presence of eukarya-like regulators for cell division, like CDC proteins, in archaea. (Buit, et al., 1996; Kawarabayasi, et al., 1998; Klenk et al., 1997). One of the CDC 48 homologues is shown to be functional in archaeal cells (Jeon et al., 1999). A Cdc6-like factor has been purified and characterized from *Sulfolobus solfataricus*. (De Felice, et al., 2003).

*Sulfolobus acidocaldarius* exhibits an unusual cell cycle, with a short prereplication period and a long post replication period. Also, in stationary phase the cell rests with two fully replicated chromosomes (Poplawski and Bernander, 1997). The tightly regulated chromosome replication, nucleoid organization, nucleoid partition and cell division have been shown to be inhibited or uncoupled by mutation (Bernander, et al., 2000). The *Sulfolobus* nucleoid shows a high level of organization at exponential phase, it is extended, arc like with clearly visible lobular structures. At stationary phase, cells are more transparent than exponential phase cells nucleoid has more relaxed organization. Nucleoid was observed to occupy more space at stationary phase as compared to exponential phase (Poplawski and Bernander., 1997).

Multiple copies of the chromosome were also observed in *Methanococcus jannaschii* (Malandrin et al., 1999), *Archaeoglobus fulgidus* and *Pyrococcus furiosus*.

**Chromosome and extra chromosomal elements of archaea:**

The archaeal chromosome is circular consisting of 1.5-4 x 10^6 bp (Klein and Schnorr, 1984). The G+C content in archaea ranges from 21 to 68 mol%. Genes are arranged in polycistronic operons. Archaeal introns have mostly been observed in tRNA and rRNA genes. Large free introns derived from pre-rRNAs have been observed as stable and abundant circular RNAs in certain Crenarchaeota (Salgia et al., 2003). Introns in
protein encoding genes are ubiquitous in eukaryotic cells, but pre-mRNA splicing is yet to be reported in archaeal and its viral genomes. Recently introns were reported in genes encoding a homolog of eukaryotic Cbf5p (centeromere binding factor 5), a subunit of a small nucleolar ribonucleoprotein, from *Aeropyrum pernix*, *Sulfolobus solfataricus* and *Sulfolobus tokodaii* (Watanabe, et al., 2002).

Plasmids isolated from hyperthermophilic archaea, Crenarchaeota (*Thermoproteales* and *Sulfolobales*) and the orders *Thermococcales* and *Thermoplasmales* in the kingdom Euryarchaeota, were found to range in size from 4.5 kb to about 40 kb. Most of them occur in high copy number. These plasmids are from relaxed to positively supercoiled in contrast to the negatively supercoiled state of DNA molecules in mesophiles, bacteria, euryarchaeae or eucaryotes. Most are cryptic, pNOB8 is conjugative, the widespread pDL10 alleviates in an unknown way autotrophic growth of its host *Desulfurolobus* by sulfur reduction. The plasmid pTIK4 appears to encode a killer function. pNOB8 has been used as a vector for the transfer of the lac z (beta-galactosidase) gene into a mutant of *S. solfataricus*. The order *Sulfolobales* has around ten plasmids. *Halobacterium salinarium* has three plasmids with covalently closed circular DNA; *Halobacterium halobium* has the plasmid, pH1. *T. acidophilum* has a 15.2 kb plasmid.

Buit et al (1996) determined the complete 1.66 mega base pair genome sequence of *Methanococcus jannaschii* and its 58 kb and 16 kb extra chromosomal elements. Only 38 % of the predicted protein encoding regions have been assigned cellular roles. Sequencing of the *Sulfolobus solfataricus* genome has shown the clustering of genes by function (Charlebois et al., 1996).

The availability of the complete genome sequence of several archaea and their analysis for frequency, location and phylogeny of archaeal mobile elements has revealed the presence of autonomous insertion sequence elements and the non-autonomous miniature inverted repeat element (MITE)-like elements. The number and diversity of elements differs greatly between the genomes. *Sulfolobus solfataricus P2* and *Halobacterium NRC-1* are very rich in elements whereas *Methanobacterium*
thermoautotrophicum contains none. The putative oriC and terC regions act as barriers for the motility of the IS and MITE -like elements in Sulfolobus solfataricus.

The genome of the archaeon Sulfolobus solfataricus contains at least four types of short sequence elements lacking open reading frames similar to the eukaryal non-autonomous mobile elements - the more conserved, SM1 and SM2 and the less conserved, SM3 and SM4, which together constitute 0.6% of the genome. (Redder et al., 2001).

Viruses of Sulfolobus have been assigned to four novel families - Fuselloviridae (viruses SSV1, SSV2 and SSV3), Rudiviridae (SIRV1 and SIRV2), Lipothrixviridae (SIFV) and Guttaviridae (SNDV) - on basis of their unique morphology. None of these viruses are lytic. Fuselloviruses are temperate whereas the other families are present in the host in a carrier state. They specifically infect thermophilic and hyperthermophilic archaea. In fusellovirus lysogens, the viral genome is integrated specifically into the host genome at a tRNA gene, by means of a virally encoded integrase, and is also present as a plasmid copy (Prangishvili et al., 2001).

**Chromatin organization in archaea:**

Chromatin structure and regulation of gene expression is determined in part by protein-DNA interactions. Histones, nuclear matrix and chromosomal scaffold proteins and enzymes involved in replication, repair and recombination contribute to chromatin structure, function and regulation.

Eukaryotic chromatin consists of DNA-protein complexes, nucleosomes where two turns of DNA are wound around an octamer of histones. The nucleosome core particle consists of 146 bp of DNA (H3+H4)2 and two H2A-H2B dimmers. Nucleosome cores are connected by a linker DNA associated with a single H1 histone molecule.

The organization of the chromatin in prokaryotes is not completely understood. The genome is present as a tightly condensed supercoiled structure, the nucleoid, within the cell (Kornberg et al., 1974; Materman and Van Gool, 1978). Nucleoid has been
shown to be associated with several DNA binding proteins, outer and inner membrane proteins, RNA polymerase core and DNA polymerases (Stonington and Pettijohn, 1971; Worcel and Burgi, 197; Portalier and Worcel, 1976; Moriya et al., 1981; Lossius et al., 1984; Yamazaki et al., 1984; Gaziev et al., 1986;). Folded chromosomes have been isolated from *Bacillus subtilis* (Guillen et al., 1978) and *E.coli* (Worcel and Burgi, 1972).

Chromatin organization in archaea has both eukaryal and eubacterial characteristics. Euryarchaea contain histones that have primary sequences in common with eukaryal nucleosome core histones. They also form a histone fold that facilitates DNA wrapping into nucleosome-like structures. Archaeal nucleosomes comprising an archaeal histone tetramer circumscribed by ~80bp of DNA, similar to the eukaryal tetrasome (H3+H4)2 with ~80bp of DNA have been reported in this organism. (Pereira et al., 1997; Sandman., et al., 1990). Archaeal nucleosomes in *Methanobacterium thermoautotrophicum* and *Methanobacterium fervidus* protect ~60 bp of DNA and multiples of 60 bp from micrococcal nuclease digestion. DNA structure which is dependent on DNA sequences directs the assembly of archaeal nucleosomes. The repeat sequence features are similar to the eukaryal nucleosome positioning elements. Repeated sequences like phased oligo (dA) tracts, 5’-(G/C)3NN (A/T)3NN-3’ and CTG repeats direct both archaeal and eukaryal nucleosome positioning (Bailey and Reeve, 1999; Sandman and Reeve, 1999). HMfB from *Methanobacterium fervidus* selectively incorporates GC, AA and TA dinucleotides at ~10 bp intervals into archaeal nucleosomes. These molecules direct the positioned assembly of archaeal nucleosomes. Fourier analysis also reveals an enrichment of the AA dinucleotide in the eucaryal and euryarchaeal genomes. Histone packaging of the genomic DNA apparently imposes constraints on genome sequence evolution (Bailey, et al, 2000).

The chromatin of *Halobacterium salinarium* consists of regions of DNA associated with protein and regions of DNA free of protein. These regions are interchangeable depending on the phase of growth (Takayanagi et al., 1992). Electron micrographs show the DNA to associate with the proteins tandemly similar to the primary structure of the eukaryotic chromosome (Shioda et al, 1989). Sheared chromosomes
obtained from the late exponential phase cells resolved into two peaks on a sucrose
density gradient. The peak I consists of protein free DNA and peak II consists of
rugged fibres consisting of nucleosome like structures. Most of the DNA is in peak I
form during exponential phase and in peak II form during stationary phase. The
transition between the two forms occurs during the late exponential phase
(Takayanagi et al., 1992).

The chromosome in *Sulfolobus acidocaldarius* is highly structured. One to two copies
of the chromosome are present during exponential phase and two copies are present
during stationary phase. The sheared chromatin of *Sulfolobus acidocaldarius* isolated
at exponential phase resolves into two peaks during gel filtration. Peak I consists of
high molecular weight aggregates of DNA and proteins. Peak II consists of high
molecular weight and an abundance of low molecular weight proteins, mainly HSNP
C’, DBNP B, HSNP A, along with proteins in the molecular weight range of 45 kDa,
25 kDa and 15 kDa (Jaya 1998). Electron microscopic studies have shown the
exponential phase nucleoid to be highly organized with a structured appearance,
occupying a relatively small part of the cell. At stationary phase the nucleoid shows a
relaxed organization, is unstructured and occupies a larger part of the cell (Poplawski
and Brenander, 1997).

**Nucleoid associated proteins in archaea:**

Four proteins, HU, H-NS, HLP1 and H have been identified as eubacterial histone-
like proteins (Drlica and Rouviere-Yaniv, 1987). HU (9 kDa) is a basic protein which
stabilizes the DNA against thermal denaturation. It wraps DNA into nucleosome like
structures having 8-10 dimers of HU and 275 bp of DNA (Rouviere-Yaniv et al.,
1979). H-NS is a neutral protein which binds intrinsically curved AT rich DNA with
slight sequence specificity (Zuber et al., 1994; Falconi et al., 1996). It does not wrap
DNA but influences the initiation of replication (Katayama et al., 1996), acts as a
transcription repressor (Ueguchi et al., 1996) and also plays a role in post
transcriptional regulation of rpo-s gene. HLP1 (17 kDa) encoded by Fir A gene
interacts with RNA polymerase and is involved in transcription and maintenance of
individual domains of supercoiling (Schaffer and Zillig, 1973; Lathe et al., 1980;
Sinden and Pettijohn, 1981). H (28 kDa) is a basic protein which binds DNA, inhibits DNA synthesis in vitro and has an amino acid sequence similar to eukaryotic histones (Hubscher et al., 1980).

Hyperthermophilic archaea grow at extreme conditions of temperature, pH and anaerobicity (Fiaia and Stetter, 1986; Bouthier et al., 1990). However their chromosomal G+C content is less than 40 mol%, *Pyrococcus* (38-45%), *Thermococcus* (38-57%) and *Methanothermus* (33 %). This demonstrates that a high G+C content is not essential for maintaining the stability of the genomic DNA in vivo at high temperatures. Many extremophiles produce low molecular weight compounds or compatible solutes to protect macromolecules especially proteins against denaturation by heat, oxidation and dryness. Compatible solutes stabilize the native tertiary structure of the protein, preventing protein aggregation. Mannosylglycerate and Mannosylglyceramide from *Rhodothermus marinus*, cyclic diphosphoglycerate from *Methanothermus fervidus*, di-my-o-inosytol phosphate from *Pyrococcus furiosus*, and diglycerol phosphate from *Archaeoglobus fulgidus* are a few examples of such compatible solutes. These solutes also protect the proteins against proteolytic degradation by stabilizing the folded state of the protein thus making the proteolytic sites less accessible to proteases (Bagyan et al., 2003).

Polyamines which are positively charged at physiological conditions stabilize the DNA against thermal denaturation by binding to the negatively recharged phosphate groups (Tomita et al., 1989; Tabor and Tabor, 1984; Morgan et al., 1987). Linear, tertiary branched and quaternary branched polyamines have been detected in hyperthermophilic archaea (Hamana, et al., 1994). High intracellular salt concentration also plays a role in maintaining double stranded DNA. *Methanococcus fervidus* and *Pyrococcus woesei* have 1 to 0.6 M potassium ion concentrations (Scholz et al., 1992; Bowater et al., 1994).

Several small basic proteins have been isolated from archaea which bind to DNA non sequence specifically. They have core domain structure similar to that of eukaryotic histones. Archaeal histones are the structural homologues of the eukaryal nucleosome core histones. Archaeal histones are not present in crenarchaeota (sulfur dependent hyperthermophiles and acidophiles) but are present in most euryarchaeota.
(methanogens, halophiles) except *Thermoplasma* (Ruepp, et al., 2000). The combination of histones of the archaeal nucleosomes varies with the growth conditions similar to the changes observed in the histone like proteins in the nucleoids of bacteria at different growth phases (Claret, et al., 1997).

HTa (9.9 kDa) from *Thermus acidophilum*, was the first archaeal histone-like protein identified (Searcy 1975, Stein and Searcy, 1978). It is a basic protein which exists as a tetramer in solution (Searcy and Stein, 1980) and protects DNA against thermal denaturation. It has a greater homology to HU than to histones H2A and H3 (Drilica and Rouviere- Yaniv, 1987; DeLange et al., 1981).

MC1 (Hmb), a 14 kDa abundant histone like protein was isolated from *Methanosarcina* species (Chartier et al., 1988; Chattier et al., 1989). It preferentially binds ds DNA (Culard et al., 1993) and protects DNA against thermal denaturation (Chartier et al., 1988) and radiolysis by fast neutrons (Isabelle et al., 1993). It stimulates transcription (Chartier et al., 1989) though it does not form nucleosomes; it preferentially binds negatively supercoiled DNA and bends the DNA upon binding (Teyssier et al., 1996). The amino acid sequence of this protein does not share homology with any of the known histones or histone like proteins.

In *Methanococcus fervidus*, proteins like HMfA and HMfB compact the DNA and protect it against thermal denaturation in vitro. The DNA-binding and nuclease-protection properties of the Hmf histones from the hyperthermophilic archaeon *Methanothermus fervidus* have been shown to be consistent with the formation of nucleosome-like structures (NLS). These proteins bind to DNA molecules as short as 20 bp and form complexes that protect DNA fragments from micrococcal nuclease (MNase) digestion that are 30 bp, approximately 60 bp and multiples of approximately 60 bp in length. DNA structure, determined by DNA sequence plays an important role in positioning and such repeated sequences as phased oligo(dA) tracts, 5'-(G/C)3NN(A/T)3NN-3' and CTG repeats direct both archaeal and eucaryal nucleosome positioning (Sandman et al., 1990; Bailey and Reeve, 1999). HpkA and HpkB from *Pyrococcus kodakaraensis* KOD1 also compact DNA and protect against thermal denaturation (Higashibata et al., 1999).
Methanobacterium formicum has three Hmf like proteins, HFoA1, HFoA2 and HFoA3. They bind and compact DNA into nucleosome like structures similar to those formed by Hmf proteins (Darcy et al., 1995).

Thermococcus species AN1 has a 4.3 kDa histone like protein, HAN1 which protects DNA against thermal denaturation (Ronimus and Musgrave, 1996a). The N-terminus of this protein shares a high degree of homology to the Hmf proteins (Ronimus and Musgrave, 1996b).

Sulfolobus genus does not have histone proteins, but it has a number of small basic and abundant DNA binding proteins ranging from 7 kDa, 8 kDa and 10 kDa which play important roles in DNA compaction, helix stabilization and DNA processing mechanisms (Kimura et al., 1984; Choli, et al., 1988; Reddy and Suryanarayana, 1988; Reddy and Suryanarayana 1989).

The 7 kDa class of proteins comprises five isoforms, 7a to 7e. The 8 kDa and 10 kDa classes of proteins comprise two proteins respectively, 8a, 8b and 10a, 10b. Electron microscopy of the protein-DNA complexes of 7 kDa, 8 kDa and 10 kDa proteins show that they bind ss and ds DNA to form compact structures (Lurz et al., 1986).

The 7d protein (known as Sso 7d, Sac7d, and Ssh7d depending on species of origin) and Alba (known as Sso 10b, Sac10b, and Ssh 10b depending on species of origin) are two highly abundant DNA binding proteins have been characterized. The non-sequence specific DNA binding protein Sso7d protects DNA against thermal denaturation and promotes the renaturation of complementary DNA strands at temperatures above the melting point of the duplex in a strictly homology dependent manner. (Guagliardi, et al., 1997). Sso7d is a 62 residue basic protein. At neutral pH it denatures at about 100 °C. Binding of multiple Sso7d molecules to short DNA fragments induces significant curvature and reduces the stiffness of the complex. It induces negative supercoiling in DNA molecules of any topology- relaxed, positively or negatively supercoiled. It compacts positively supercoiled and relaxed forms of DNA but not negatively supercoiled DNA. It also inhibits the positive supercoiling activity of the thermophile specific enzyme, reverse gyrase (Napoli, et al., 2002).
Family B DNA polymerase (polB1) from the hyperthermophile *Sulfolobus solfataricus* degrades both ss and ds DNA at similar rates in vitro at physiological temperatures. Studies on the effect of the 7 kDa proteins on the polymerization and 3′-5′ exonuclease activities of this enzyme showed that they modulate the extension and excision activities thus reducing the cost of proofreading. The 7 kDa proteins inhibit the excision and enhance the extension of matched template primers. They however do not protect the ss DNA from cleavage by polB1. They also do not affect the proof reading ability of polB1 and do not inhibit the excision of mismatched primers. The dNTP concentrations required for effective inhibition of 3′-5′ exonuclease activity of polB1 was lowered from 1mM to 50 μM in the presence of the 7 kDa proteins at 65°C (Lou, et al., 2004). So far the 7 kDa proteins have only been observed in *Sulfolobus* species, wherein it is highly expressed and highly conserved.

Archaeal proteins of Sac 10b family are thought to affect the topology of the chromosomal DNA in thermophilic archaea. Ssh 10b constrains negative DNA supercoils in a temperature dependent manner (Xue et al., 2000). Studies have shown that two forms of Ssh 10b homodimers coexist in solution and the slow cis-trans isomerization of the Leu61-Pro62 peptide bond accounts for the conformational heterogeneity of the homodimer. The trans form (T-form) is dominant at high temperatures and the cis form (C-form) at lower temperatures. Both forms have the same DNA binding site but different conformational features which are responsible for the temperature dependent nature of Ssh 10b-DNA interaction. (Cui, et al., 2003). The alba proteins coat the ds DNA without significantly compacting it protecting it from nuclease digestion (Lurz et al., 1986). Recent reports show that Ssh 10b binds with similar affinity to ds DNA, ss DNA and RNA invitro and binds exclusively to RNA in *Sulfolobus shibitae* cells invivo (Guo et al., 2003).

Alba where present, always occurs along with a second chromatin protein, a histone, 7d protein or HU homologue. Archaeal species lacking alba encode histones along with another DNA binding protein, such as MC1 in mesophilic methanogens. This leads to the speculation of interaction or cooperation between more than one type of chromatin proteins to compact DNA fully in archaea (White and Bell, 2002).
Reddy and Suryanarayana, (1988) purified four acid soluble low molecular weight proteins from the nucleoid of Sulfolobus acidocaldarius, HSNP (Helix stabilizing nucleoid proteins) A (12 kDa), C (9 kDa) and C (7 kDa) and DBNP (DNA binding nucleoid protein) B (10 kDa). The HSNP proteins stabilize the DNA helix and protect DNA against thermal denaturation. They are present exclusively on the genomic DNA as revealed by immuno gold labeling electron microscopy of the nucleoid and could be responsible for DNA compaction and genome organization (Bohrmann et al., 1994). HSNP C is identical to Sac 7d (Gauri 1997) and DBNP B is similar to Sac 10b (Sreenivas 1994). DBNPB binds ds and ss DNA strongly but does not protect against thermal denaturation. It forms different types of novel complexes with DNA at different protein to DNA ratios (Sreenivas et al., 1998).

HSNP A strongly binds native DNA and RNA, weakly to denatured DNA and ribosomes. It protects the DNA against thermal denaturation but not RNA.

**Growth phase dependent variation in content and levels of proteins associated with nucleoid:**

Structural and functional modulations occur in the nucleoid of the prokaryotes and chromatin of eukaryotes during the transition from exponential growth to stationary phase. Cells enter into stationary phase upon sensing an impending saturation level of their population density.

Nucleoid proteins play functional roles besides their structural roles, in the regulation of essential DNA functions such as replication, recombination and transcription. Studies on the composition of these nucleoid associated proteins show that a variation in protein contents during the stationary phase is accompanied by compaction of the genome DNA and silencing of the genome functions.

In *E.coli*, the major DNA binding proteins, Fis, Hu and Hfq during exponential phase are replaced by a single protein, Dps at stationary phase. CbpA, the curved DNA binding protein appears only in the late stationary phase (Ishihama, 1999, Azam et al., 1999).
The observed increase in density of cells at stationary phase has been reported to require the presence of the RNA polymerase RpoS (sigma-S) needed for stationary phase gene transcription (Makinoshima, et al., 2003).

Growth phase dependent expression and degradation of histones has been observed in the thermophilic archaeon Thermococcus zilligii, the levels of certain proteins decreased dramatically on entry into stationary phase. The HTz protein could not be detected by late stationary phase (Dinger et al., 2000).

*Methanothermus fervidus* cells have been shown to contain HmfA and HmfB homodimers and HmfA-HmfB heterodimers. The HmfA/HmfB ratio and the relative amounts of homodimers and heterodimers is growth phase regulated. HmfA is more abundant in exponential phase and HmfB during the stationary phase (Sandman, 1994).

In the present study detailed characterization of *S. acidocaldarius* nucleoid was undertaken to understand the organization of the intracellular DNA in this Crenarchaeote. Studies were also done on HSNP A to characterize its biochemical properties and analyze its DNA binding properties.
Objectives and scope of the present investigation

Intracellular DNA in eukaryotes is organized into chromatin with nucleosomal repeat structure formed by the interaction of DNA with basic proteins, histones. In contrast, in eubacteria like *E. coli* the DNA is in condensed state with no clear defined structural organization. In eubacteria, several histone-like proteins have been found associated with intracellular DNA with no nucleosomal organization.

Archaea constitute a separate domain of organisms sharing some properties with eukarya and eubacteria and some unique properties of their own. In euryarchaeota the primitive nucleosomal organization is seen in case of *M. fervidus* which has histone-like homologs, Hmf A and Hmf B which interact with DNA forming nucleosomal structure with 80 bp DNA repeat structure. Not much work has been carried out on the chromatin organization in crenarchaeota.

In the present investigation, work has been carried out on the isolation and biochemical characterization of the nucleoid from the crenarchaeote *S. acidocaldarius*, an organism which survives at extreme conditions of temperature (75-80 °C), pH (3.0), has a low G+C content (40%) and also lacks a cell wall.

The composition of nucleic acid and protein from the nucleoid was analyzed. The components involved in the structural organization of the nucleoid were determined and analyzed by employing various biochemical techniques. The results indicate that the nucleoid of *S. acidocaldarius* is similar to eubacteria. The nucleoid was found to occur in two forms, membrane associated form and membrane free nucleoid.

Important results of the present investigation are detection of S-layer protein in the membrane associated nucleoid, observance of abundant RNA which helps formation of nucleic acid protein aggregates and variations observed in the nucleoid organization and composition during stationary phase of growth.

Work has also been carried out on the solution state and DNA binding properties of nucleoid associated protein HSNP A. The results indicate that the protein binds both
single stranded and double stranded DNA. It interacts by coating the DNA strands starting from the free ends.

Further work is necessary to understand mode of interaction of S-layer proteins with nucleoid DNA and the role of RNA and other protein components in the condensed nucleoid structure. Analysis of the interaction between HSNP A and HSNP C and the subsequent role in nucleoid condensation in vivo needs to be worked out.